

Weekly Report 08/03/2014

Visual Data Inspection Project (VIP)

I've been redesigning view layouts for the visual data inspection project. The new layout would be based on node-link (or view-link to be more specific). Basic interactions are to select new centric view and generate related views, just like the large-single-small-multiples approach. Additional interactions would include view merging/collapsing, view annotating, view grouping, view pinning and view deleting.

Implementation is in slow progress since I have to do implementing new features and getting to know existing codes in parallel. I will speed up when I'm familiar with the codes.

Paper Reading

Paper: Small multiples, large singles: A new approach for visual data exploration [Elzen2013]

This paper basically does the same thing with our VIP project: visual data exploration using small multiples, or coordinated views.

Their work adopts the filmstrip metaphor to display the exploration process. Each visible display (or frame) of the filmstrip is one large single with several small multiples generated according to the large single. Generation of small multiples can be based on mappings of other visualization, or filtering of data distribution, or parameters in generating layouts (e.g. number of clusters). Users explore a dataset by keeping selecting new large singles and thus generating new small multiples. Users can also revisit the exploration process by going back and forth to other frames of the filmstrip.

The evaluation they did is based on the comparison between their small multiples system and a dual view mechanism on visual exploration functions including identification, correlation, comparison, clustering and exploration. A user study is conducted and evaluated for this comparison.

Discussion: This work is better than our first version in all aspects. The filmstrip metaphor with large singles and small multiples is clever to address the problem of utilizing display space while maintain exploration process. But I doubt its efficiency in making connections of views. To continue our work, I will switch our focus to view layout and view interactions (see appendix at the end).

[Elzen2013] van den Elzen, Stef, and Jarke J. van Wijk. "Small multiples, large singles: A new approach for visual data exploration." *Computer Graphics Forum*. Vol. 32. No. 3pt2. Blackwell Publishing Ltd, 2013.

Miscellaneous

House moving: really a tough work

TODO

VAST Challenge 2-page summary writing.

Contributions and Task Descriptions of Data Inspection

Motivation/Goal

To provide users a tool for more awareness of a new dataset

Contribution

- A comprehensive interactive data inspection technique that reflects both users' cognitive process and level of awareness towards data;
Cognitive process is represented by iteratively developing of view-link diagram.
Awareness includes statistical distributions and correlations of different dimensions of data.
- Multi-facet displays of data that indicate either one or combinational dimensions of data;
Views include single and combinational dimensions, various statistical distributions of data.
- View-based interactions that conduct intuitive reasoning behaviors.
Interactions include view elaborating (view creating and view merging), view organizing (view deleting, view grouping and view collapsing), view annotating and view pinning (position-fixing).

Task Description

Data Diagnosis Tasks

- **Task 1** - How many blank values are there in each dimension, and are they correlated?
- **Task 2** - Is the data continuous or not continuous in terms of time?
- **Task 3** - Is the data reasonable or unreasonable in terms of the context of data?

Data Inspection Tasks

- **Task 4** - How are the values in each dimension and combination of dimensions distributed?
- **Task 5** - Are there any redundant dimensions?
Normal dataset usually contains redundant dimensions, such as dimensions like group and subgroup, which would draw too much users' attention. To efficiently utilize users'
- **Task 6** - Compare values in terms of aggregated time.
- **Task 7** - Explore more insights out of data.